# Bayesian analysis of Non-Homogeneous Poisson Processes

**Evans Gouno**
Université de Bretagne Sud
Campus de Tohannic
Vannes – FRANCE
evans.gouno@univ-ubs.fr

## Abstract

Many reliability studies involve counting process like Poisson process. A problem arising using this model is the question of the nature of the failure intensity. In particular, one would like to assess whether this latter is constant (homogeneous Poisson process) or increasing (non homogeneous Poisson process). The purpose of this work is to propose a Bayesian approach to test failure intensity associated to Poisson process on the basis of data provided according the following scheme. We consider a number of failure dates in a time window split in equally space intervals of a certain length which is a parameter to be estimated. The failure rate is considered to be a step function that is to say it is constant over the equally spaced intervals. Reasoning on the interarrival times, the likelihood is obtained and Gamma Prior distributions for the different values of the failure rate which are supposed to be independent, are assumed to derive Bayesian estimation. The Bayes factor is computed to test hypothesis of a constant rate model against a non constant rate model. A simulation study is conducted to investigate the efficiency of the procedure.

## 1 Introduction

Non-homogeneous poisson processes are commonly used to model reliability growth of complex systems or reliability of repairable systems. When the intensity function of the process is of the form : $\lambda(t) = \beta t^{\beta-1}/\alpha^{\beta}$, the process is called a Weibull process or power-law process. This process has been widely studied. Finkelstein (1976) proposed confidences intervals for $\alpha$ and $\beta$. Goodness-of-fit tests are suggested by Park and Kim (1992). Park and Pickering (1997) focuse on the shape parameter inference testing HPP versus NHPP. Some diffrent form of intensity function has been also suggested and studied in particular for software reliability analysis. Refer to Hosain, Dahiya (1993) for a quick review. Guida and al. (1989) consider Bayes inference for the Weibull process and compare Bayes estimation and maximum likelihood estimation. Raftery and Akman (1986) describe a bayesian analysis of a Poisson process with a change-point. We consider here making inference on NHPP depicted as a Poisson process with several change-point leading to a semi-parametric approach of the inference problem. We assume a non-homogeneous Poisson process with a failure intensity $\lambda(t)$ such as $\lambda(t) = \lambda_i$ for $t \in [(i-1)\tau, i\tau]$ . The analysis relies on $n$ events which occur at times $t = (t_1, \cdots, t_n)$ on an observation period $[0, C]$, $(C = m\tau)$. In section 2, we derive bayes estimators for $\lambda = (\lambda_1, \cdots, \lambda_m)$ and $\tau$. In section 3, we propose a test to decide between $\mathcal{M}_0$ a constant failure intensity model and $\mathcal{M}_1$ the model initially considered that is to say a non-homogeneous Poisson process. To end with, we study the procedure on simulated data in section 4.

## 2 Estimation

Let $t = (t_1, \cdots, t_n)$ be $n$ dates of events. The likelihood is :

$$f(t|\lambda) = \prod_{i=1}^{m} \lambda_i^{N_i(\tau)} e^{-\lambda_i \tau}$$

with $N_i(\tau)$ , the number of events in $[(i-1)\tau, i\tau]$. We assume that $\lambda_1, \lambda_2, \cdots, \lambda_m$ are independent a priori, and that they have a Gamma distribution:

$$\pi(\lambda_i) \propto \lambda_i^{\alpha_i - 1} e^{-\beta_i \lambda_i} \ , \quad i = 1, \cdots, m.$$

The joint posterior distribution is then :

$$\pi(\lambda, \tau | t) = \prod_{i=1}^{m} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda_i^{\alpha_i + N_i(\tau) - 1} e^{-(\tau + \beta_i)\lambda_i} \pi(\tau) / f(t) \tag{1}$$

where $\pi(\tau)$ is the prior distribution of $\tau$ and $f(t)$ the predictive that is to say:

$$\int f(t|\lambda, \tau) \pi(\lambda) \pi(\tau) d\lambda_1 \cdots d\lambda_m d\tau. \tag{2}$$

Let

$$\varphi_m(\tau) = \prod_{j=1}^{m} \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \frac{\Gamma(N_j(\tau) + \alpha_j)}{(\tau + \beta_j)^{\alpha_j + N_j(\tau)}} \quad \text{and} \quad \varphi_{m,i}(\tau) = \prod_{j=1, i \neq j}^{m} \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \frac{\Gamma(N_j(\tau) + \alpha_j)}{(\tau + \beta_j)^{\alpha_j + N_j(\tau)}}$$

With these notations, we have $f(t) = \int_0^{+\infty} \varphi_m(\tau) \pi(\tau) d\tau$.
The marginal distribution for $\tau$ is obtained by integrating (1) with respect to $(\lambda_1, \cdots, \lambda_m)$, we have:

$$\pi(\tau|t) = \varphi_m(\tau) \pi(\tau) / f(t)$$

and the marginal for $\lambda_i$ is obtained by integrating (1) with respect to $\tau$ and to $\lambda_j$, $j = 1, \cdots, i-1, i+1, \cdots, m$:

$$\pi(\lambda_i|t) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \int_0^{+\infty} \lambda_i^{\alpha_i + N_i(\tau) - 1} e^{-(\tau + \beta_i)\lambda_i} \varphi_{m,i}(\tau) \pi(\tau) d\tau / f(t) \ , \quad i = 1, \cdots, m.$$

If a quadratic loss function is considered, the bayes estimator $\tilde{\tau}$ and $\tilde{\lambda}$ for $\tau$ and $\lambda$ are given by the posterior expectation :

$$\tilde{\tau} = \int_0^{+\infty} \tau \varphi_m(\tau) \pi(\tau) / f(t) \quad \text{and} \quad \tilde{\lambda}_i = \int_0^{+\infty} \frac{N_i(\tau) + \alpha_i}{\tau + \beta_i} \varphi_m(\tau) \pi(\tau) d\tau / f(t) \ , \quad i = 1, \cdots, m.$$

We describe in section 4, methods to compute these integrals and thus estimates for $\tau$ and $\lambda$ .

## 3 Hypothesis Testing

Let $\mathcal{M}_0$ be the model with a constant failure intensity and $\mathcal{M}_m$ the model with a non constant failure intensity. To decide between the two models, we compute the Bayes factor :

$$B_{(0,m)} = \frac{f(t|\mathcal{M}_0)}{f(t|\mathcal{M}_m)}.$$

For the model $\mathcal{M}_0$, we assume a gamma prior distribution with parameters $(\alpha_0, \beta_0)$. The predictive is then:

$$f(t|\mathcal{M}_0) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} \frac{\Gamma(n + \alpha_0)}{(t_n + \beta_0)^{n + \alpha_0}}.$$

$f(t|\mathcal{M}_1)$ is $f(t)$ defined previously (2).

This approach can be extended in the following way. Let $\mathcal{M}_q$ be a model with a $q$ steps ($q-1$ change-points) failure intensity. The predictive is : $f(t|\mathcal{M}_q) = \int_0^{+\infty} \varphi_q(\tau) \pi(\tau) d\tau$. Thus computing the bayes factor $B_{(m,q)}$ that is to say $f(t|\mathcal{M}_m)/f(t|\mathcal{M}_q)$ allows to test the number of steps and to decide between $\mathcal{M}_m$ and $\mathcal{M}_q$.

# 4 Application

To compute $\tilde{\tau}$ and $\tilde{\lambda}_i$, $i = 1, \cdots, m$, we use the Monte-Carlo method to approximate integral.
$f(t|\mathcal{M}_m)$ is approximated by $\sum_{l=1}^{M} \varphi_m(\tau_l^*)/M$ with $\tau_l^*$ a realisation of $\tau$ according to $\pi(\tau)$, $\tilde{\tau}$ by $\frac{1}{M}\sum_{l=1}^{M} \tau_l^* \varphi_m(\tau_l^*)/f(t)$ and $\tilde{\lambda}_i$ by $\frac{1}{M}\sum_{l=1}^{M} \frac{N_i(\tau_l^*)+\alpha_i}{\tau_l^*+\beta_i}\varphi_m(\tau_l^*)/f(t)$, $i = 1, \cdots, m$.
We simulate dates corresponding to a failure intensity with three steps : $\lambda_i = i/10$, $i = 1, 2, 3$ with $\tau = 50$. To construct prior we deduce values for $(\alpha, \beta)$ approximating $\lambda$ and considering $\lambda \sim \alpha/\beta$. We choose $\alpha_i$ in order of magnitude of $N_i(\tau)$ and we choose $\beta_i$ in order of magnitude of $\tau$. We apply the method with three types of prior parameters for $\lambda$ which are displayed in the following table :

Table 1: Prior values parameters on $\lambda$.

| Prior on $\lambda$ | $(\alpha_1, \beta_1)$ | $(\alpha_2, \beta_2)$ | $(\alpha_3, \beta_3)$ |
|---|---|---|---|
| $\mathcal{P}_1$ | (2.25 , 15) | (6.25 , 25) | (12.25 , 35) |
| $\mathcal{P}_2$ | (1.6 , 40) | (1.6 , 40) | (1.6 , 40) |
| $\mathcal{P}_3$ | (3 , 60) | (2, 60) | (1, 60) |

A normal distribution is considered for $\tau$ with parameter $(\mu, \sigma^2)$.
The following table gives the results:

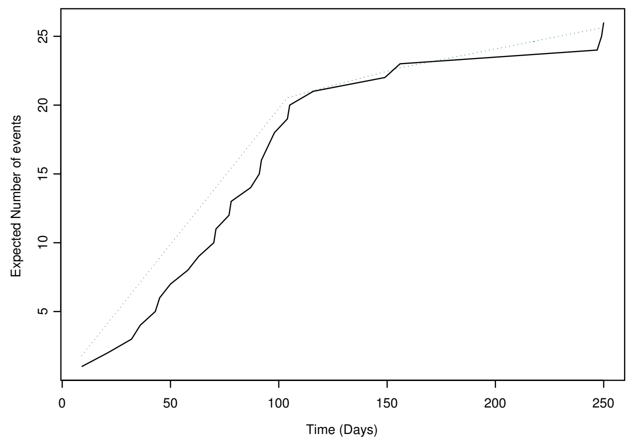
Table 2: Bayes estimates.

| Prior on $\lambda$ | Prior on $\tau$ | | Bayesian estimates | | | |
|---|---|---|---|---|---|---|
| $(\alpha, \beta)$ | $\mu$ | $\sigma^2$ | $\tilde{\lambda}_1$ | $\tilde{\lambda}_2$ | $\tilde{\lambda}_3$ | $\tilde{\tau}$ |
| | 20 | 5 | 0.19995 | 0.19997 | 0.25850 | 16.25 |
| | 40 | 5 | 0.15556 | 0.16204 | 0.31790 | 44.42 |
| $\mathcal{P}_1$ | 50 | 5 | 0.13815 | 0.18517 | 0.31338 | 51.95 |
| | 50 | 10 | 0.13775 | 0.19251 | 0.25686 | 59.40 |
| | 60 | 5 | 0.13637 | 0.18309 | 0.31025 | 52.82 |
| | 20 | 5 | 0.12625 | 0.04319 | 0.04319 | 20.19 |
| | 40 | 5 | 0.10505 | 0.04976 | 0.10505 | 32.34 |
| $\mathcal{P}_2$ | 50 | 5 | 0.10461 | 0.08028 | 0.17759 | 42.20 |
| | 50 | 10 | 0.09636 | 0.09636 | 0.17480 | 49.24 |
| | 60 | 5 | 0.09258 | 0.17937 | 0.06364 | 63.69 |
| | 20 | 5 | 0.10156 | 0.08886 | 0.08886 | 18.76 |
| | 40 | 5 | 0.09915 | 0.09915 | 0.14873 | 40.85 |
| $\mathcal{P}_3$ | 50 | 5 | 0.09349 | 0.10284 | 0.19634 | 46.95 |
| | 50 | 10 | 0.09140 | 0.10969 | 0.20109 | 49.39 |
| | 60 | 5 | 0.09402 | 0.16455 | 0.09402 | 67.62 |

With $\mathcal{P}_1$ which corresponds to prior close to imput values (close in the sense of the distance $\sum_i(\lambda_i - \tilde{\lambda}_i)^2$) is obtained with $\mu = 60$ and $\sigma^2 = 5$. Remark that $\mu = 20$ that is "misleading" on prior steps length, leads to better estimates than estimates obtained with $\mathcal{P}_2$ and $\mathcal{P}_3$. Using $\mathcal{P}_2$ suppose a constant prior failure intensity and the best estimate in this case are obtained with $\mu$=50 and $\sigma^2$=10. With $\mathcal{P}_3$, which corresponds to a prior rather far from the input data the estimation results are not satisfying like with $\mathcal{P}_2$.
We consider now data from the Naval Fleet Computer Programming Center, Naval Tactical Data System (NTDS). These data are displayed by Hossain and Dayiha (1993) and used by many authors for model validation. They concern a software developpment. 26 dates of failures (in days) are recorded. The first failure occurs after 9 days, the last failure occurs at 250 days. We apply the previous method to assess a three-step failure intensity. We find $\tilde{\tau}$=112,38 days , $\tilde{\lambda}_1$=0,1975, $\tilde{\lambda}_2$=0,04158 and $\tilde{\lambda}_3$= 0.03118 with prior $\tau \sim \mathcal{N}(100, 10)$, and $(\alpha, \beta)$= ((18,80),(5,80) (3,80)).
Computation of the bayes factor to test constant failure intensity with prior $(\alpha_0, \beta_0) = (10, 100)$ against three-step failure intensity gives the following value: $B_{(0,3)}$= 0.126. Thus we conclude that we cannot reject

that the NHPP model is adequate for the NTDS data. The following figure represents the observed expected number of events and the fitted Bayes model (dash line).

Figure 1 : Plots of the Expected Number of Events.




## 5 Concluding remarks

This work gives the basis for a Bayesian semi-parametric inference on Poisson process failure intensity. We demonstrate with a somewhat very simple example the feasabilty of the approach. It appears to be a more flexible alternative to the classical parametric approach. More works need to be done in particular to study the efficiency of the method through simulations. Some hints concerning the choice of prior parameters have been given allowing practising statisticians to apply the method. The problem of testing the nature of the failure intensity is interpreted in term of testing number of steps. The Bayes factor computation gives a convenient way to handle this question. A further work direction is the case where steps lengths are different.

## References


Finkelstein J. M., (1976). Confidence bounds on the Parameters of the Weibull Process. *Technometrics 18*(1),115–117.

Guida M., Calabria R., Pulcini G. (1989). Bayesian Inference for Non-Homogeneous Poisson Process with Power Intensity Law. *IEEE Transactions on Reliability, 38*(5), 603–609.

Hossain S. A., Dahiya R. C. (1993). Estimating the Parameters of a Non-Homogeneous Poisson-Process Model for Software Reliability. *IEEE Transactions on Reliability, 42*(4), 604–612.

Park W. J., Kim Y. G. (1992). Goodness-of-Fit Tests for the Power-Law Process. *IEEE Tansaction on Reliability 41*(1),107–111.

Park W. J., Pickering E. H. (1997). Statistical Analysis of a Power-Law Model for Repair Data. *IEEE Tansaction on Reliability 46*(1), 27–30.

Raftery A. E., Akman V. E. (1986). Bayesian analysis of a Poisson process with a change-point. *Biometrika 73*(1),85 –9.